

K.
Kerr
PAGE: 1

SHK

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/302,024

DATE: 04/05/2000
TIME: 17:05:33

INPUT SET: S35223.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING

ENTERED

1
2
3 (1) General Information:
4
5 (i) APPLICANT:
6 (A) NAME: Hodgson, John
7 (B) STREET: Great West Road
8 (C) CITY: Brentford
9 (D) STATE OR PROVINCE: Middlesex
10 (E) COUNTRY: United Kingdom
11 (F) POSTAL CODE: TW8 9EP
12
13
14 (A) NAME: Hodgson, John
15 (B) STREET: Great West Road
16 (C) CITY: Brentford
17 (D) STATE OR PROVINCE: Middlesex
18 (E) COUNTRY: United Kingdom
19 (F) POSTAL CODE: TW8 9EP
20
21 (ii) TITLE OF INVENTION: Novel tRNA Synthetase
22
23
24 (iii) NUMBER OF SEQUENCES: 2
25
26 (v) COMPUTER READABLE FORM:
27 (A) MEDIUM TYPE: Diskette
28 (B) COMPUTER: IBM Compatible
29 (C) OPERATING SYSTEM: DOS
30 (D) SOFTWARE: FastSEQ for Windows Version 2.0
31
32 (vi) CURRENT APPLICATION DATA:
33 (A) APPLICATION NUMBER: 09/302,024
34 (B) FILING DATE:
35
36 (vii) PRIOR APPLICATION DATA:
37 (A) APPLICATION NUMBER: US/08/785,455
38 (B) FILING DATE:
39
40 (2) INFORMATION FOR SEQ ID NO:1:
41
42 (i) SEQUENCE CHARACTERISTICS:
43 (A) LENGTH: 1974 base pairs
44 (B) TYPE: nucleic acid
45 (C) STRANDEDNESS: double
46 (D) TOPOLOGY: linear

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/302,024DATE: 04/05/2000
TIME: 17:05:34

INPUT SET: S35223.raw

47
48 (ii) MOLECULE TYPE: Genomic DNA
49
50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
51
52 ATGGCTAAAG AAACATTTTA TATAACAACC CCAATATACT ATCCTAGTGG GAATTTACAT 60
53 ATAGGACATG CATATTCTAC AGTGGCTGGA GATGTTATTG CAAGATATAA GAGAATGCAA 120
54 GGATATGATG TTCGTTATTT GACTGGAACG GATGAACACG GTCAAAAAAT TCAAGAAAAA 180
55 GCTCAAAAAG CTGGTAAGAC AGAAATTGAA TATTTGGATG AGATGATTGC TGGAAATTA 240
56 CAATTGTGGG CTAAGCTTGA AATTTCAAAT GATGATTTTA TCAGAACAAC TGAAGAACGT 300
57 CATAAACATG TCGTTGAGCA AGTGTTTGAA CGTTTATTAA AGCAAGGTGA TATCTATTTA 360
58 GGTGAATATG AAGGTTGGTA TTCTGTTCCT GATGAAACAT ACTATACAGA GTCACAATTA 420
59 GTAGACCCAC AATACGAAAA CGGTAAAAAT ATTGGTGGCA AAAGTCCAGA TTCTGGACAC 480
60 GAAGTTGAAC TAGTTAAAGA AGAAAGTTAT TTCTTTAATA TTAGTAAATA TACAGACCGT 540
61 TTATTAGAGT TCTATGACCA AAATCCAGAT TTTATACAAC CACCATCAAG AAAAAATGAA 600
62 ATGATTAACA ACTTCATTAA ACCAGGACTT GCTGATTTAG CTGTTTCTCG TACATCATTT 660
63 AACTGGGGTG TCCATGTTCC GTCTAATCCA AAACATGTTG TTTATGTTTG GATTGATGCG 720
64 TTAGTTAACT ATATTTCAGC ATTAGGCTAT TTATCAGATG ATGAGTCACT ATTTAACAAA 780
65 TACTGGCCAG CAGATATTCA TTTAATGGCT AAGGAAATTG TGCGATTCCA CTCAATTATT 840
66 TGGCCTATTT TATTGATGGC ATTAGACTTA CCGTTACCTA AAAAAGTCTT TGCACATGGT 900
67 TGGATTTTGA TGAAAGATGG AAAAATGAGT AAATCTAAAG GTAATGTTGT AGACCCTAAT 960
68 ATTTTAATTG ATCGCTATGG TTTAGATGCT ACACGTTATT ATCTAATGCG TGAATTACCA 1020
69 TTTGGTTCAG ATGGCGTATT TACACCTGAA GCATTTGTTG AGCGTACAAA TTTTCGATCTA 1080
70 GCAAATGACT TAGGTAACTT AGTAAACCGT ACGATTTCTA TGGTTAATAA GTACTTTGAT 1140
71 GGCGAATTAC CAGCGTATCA AGGTCCACTT CATGAATTAG ATGAAGAAAT GGAAGCTATG 1200
72 GCTTTAGAAA CAGTGAAAAG CTACACTGAA AGCATGGAAA GTTTGCAATT TTCTGTGGCA 1260
73 TTATCTACGG TATGGAAGTT TATAAGTAGA ACGAATAAGT ATATTGACGA AACAACGCCT 1320
74 TGGGTATTAG CTAAGGACGA TAGCCAAAAA GATATGTTAG GCAATGTAAT GGCTCACTTA 1380
75 GTTGAAAATA TTCGTTATGC AGCTGTATTA TTACGTCCAT TCTTAACACA TGCGCCGAAA 1440
76 GAGATTTTTG AACAATTGAA CATAAACAAT CCTCAATTTA TGGAATTTAG TAGTTTAGAG 1500
77 CAATATGGTG TGCTTACTGA GTCAATTATG GTTACTGGGC AACCTAAACC TATTTTCCCA 1560
78 AGATTGGATA GCGAAGCGGA AATTGCATAT ATCAAAGAAT CAATGCAACC GCCTGCTACT 1620
79 GAAGAGGAAA AAGAAGAGAT TCCTAGCAAA CCTCAAATTG ATATTAAAGA CTTTGATAAA 1680
80 GTTGAAATTA AGGCAGCAAC GATTATTGAT GCTGAACATG TTAAGAAGTC AGATAAGCTT 1740
81 TTAAAAATTC AAGTAGACTT AGATTCTGAA CAAAGACAAA TTGTATCAGG AATTGCCAAA 1800
82 TTCTATACAC CAGATGATAT TATTGGTAAA AAAGTAGCAG TTGTTACTAA CCTGAAACCA 1860
83 GCTAAATTAA TGGGACAAAA ATCTGAAGGT ATGATATTAT CTGCTGAAAA AGATGGTGTA 1920
84 TTAACCTTAG TAAGTTTACC AAGTGCAATT CCAAATGGTG CAGTGATTAA ATAA 1974
85
86 (2) INFORMATION FOR SEQ ID NO:2:
87
88 (i) SEQUENCE CHARACTERISTICS:
89 (A) LENGTH: 657 amino acids
90 (B) TYPE: amino acid
91 (C) STRANDEDNESS: single
92 (D) TOPOLOGY: linear
93
94 (ii) MOLECULE TYPE: peptide
95
96 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
97
98 Met Ala Lys Glu Thr Phe Tyr Ile Thr Thr Pro Ile Tyr Tyr Pro Ser
99 1 5 10 15

RAW SEQUENCE LISTING PATENT APPLICATION US/09/302,024

DATE: 04/05/2000
TIME: 17:05:34

INPUT SET: S35223.raw

100	Gly	Asn	Leu	His	Ile	Gly	His	Ala	Tyr	Ser	Thr	Val	Ala	Gly	Asp	Val
101				20					25					30		
102	Ile	Ala	Arg	Tyr	Lys	Arg	Met	Gln	Gly	Tyr	Asp	Val	Arg	Tyr	Leu	Thr
103			35					40					45			
104	Gly	Thr	Asp	Glu	His	Gly	Gln	Lys	Ile	Gln	Glu	Lys	Ala	Gln	Lys	Ala
105		50					55					60				
106	Gly	Lys	Thr	Glu	Ile	Glu	Tyr	Leu	Asp	Glu	Met	Ile	Ala	Gly	Ile	Lys
107	65				70						75				80	
108	Gln	Leu	Trp	Ala	Lys	Leu	Glu	Ile	Ser	Asn	Asp	Asp	Phe	Ile	Arg	Thr
109				85					90						95	
110	Thr	Glu	Glu	Arg	His	Lys	His	Val	Val	Glu	Gln	Val	Phe	Glu	Arg	Leu
111				100					105					110		
112	Leu	Lys	Gln	Gly	Asp	Ile	Tyr	Leu	Gly	Glu	Tyr	Glu	Gly	Trp	Tyr	Ser
113			115					120					125			
114	Val	Pro	Asp	Glu	Thr	Tyr	Tyr	Thr	Glu	Ser	Gln	Leu	Val	Asp	Pro	Gln
115		130					135					140				
116	Tyr	Glu	Asn	Gly	Lys	Ile	Ile	Gly	Gly	Lys	Ser	Pro	Asp	Ser	Gly	His
117	145				150					155					160	
118	Glu	Val	Glu	Leu	Val	Lys	Glu	Glu	Ser	Tyr	Phe	Phe	Asn	Ile	Ser	Lys
119				165					170					175		
120	Tyr	Thr	Asp	Arg	Leu	Leu	Glu	Phe	Tyr	Asp	Gln	Asn	Pro	Asp	Phe	Ile
121			180					185						190		
122	Gln	Pro	Pro	Ser	Arg	Lys	Asn	Glu	Met	Ile	Asn	Asn	Phe	Ile	Lys	Pro
123			195				200					205				
124	Gly	Leu	Ala	Asp	Leu	Ala	Val	Ser	Arg	Thr	Ser	Phe	Asn	Trp	Gly	Val
125		210					215					220				
126	His	Val	Pro	Ser	Asn	Pro	Lys	His	Val	Val	Tyr	Val	Trp	Ile	Asp	Ala
127	225				230						235				240	
128	Leu	Val	Asn	Tyr	Ile	Ser	Ala	Leu	Gly	Tyr	Leu	Ser	Asp	Asp	Glu	Ser
129				245					250					255		
130	Leu	Phe	Asn	Lys	Tyr	Trp	Pro	Ala	Asp	Ile	His	Leu	Met	Ala	Lys	Glu
131			260					265						270		
132	Ile	Val	Arg	Phe	His	Ser	Ile	Ile	Trp	Pro	Ile	Leu	Leu	Met	Ala	Leu
133		275					280						285			
134	Asp	Leu	Pro	Leu	Pro	Lys	Lys	Val	Phe	Ala	His	Gly	Trp	Ile	Leu	Met
135		290				295					300					
136	Lys	Asp	Gly	Lys	Met	Ser	Lys	Ser	Lys	Gly	Asn	Val	Val	Asp	Pro	Asn
137	305				310					315					320	
138	Ile	Leu	Ile	Asp	Arg	Tyr	Gly	Leu	Asp	Ala	Thr	Arg	Tyr	Tyr	Leu	Met
139				325					330					335		
140	Arg	Glu	Leu	Pro	Phe	Gly	Ser	Asp	Gly	Val	Phe	Thr	Pro	Glu	Ala	Phe
141			340					345					350			
142	Val	Glu	Arg	Thr	Asn	Phe	Asp	Leu	Ala	Asn	Asp	Leu	Gly	Asn	Leu	Val
143			355				360					365				
144	Asn	Arg	Thr	Ile	Ser	Met	Val	Asn	Lys	Tyr	Phe	Asp	Gly	Glu	Leu	Pro
145		370				375					380					
146	Ala	Tyr	Gln	Gly	Pro	Leu	His	Glu	Leu	Asp	Glu	Glu	Met	Glu	Ala	Met
147	385				390					395					400	
148	Ala	Leu	Glu	Thr	Val	Lys	Ser	Tyr	Thr	Glu	Ser	Met	Glu	Ser	Leu	Gln
149				405					410					415		
150	Phe	Ser	Val	Ala	Leu	Ser	Thr	Val	Trp	Lys	Phe	Ile	Ser	Arg	Thr	Asn
151			420					425				430				
152	Lys	Tyr	Ile	Asp	Glu	Thr	Thr	Pro	Trp	Val	Leu	Ala	Lys	Asp	Asp	Ser

RAW SEQUENCE LISTING PATENT APPLICATION US/09/302,024

DATE: 04/05/2000

TIME: 17:05:34

INPUT SET: S35223.raw

```

153          435          440          445
154 Gln Lys Asp Met Leu Gly Asn Val Met Ala His Leu Val Glu Asn Ile
155          450          455          460
156 Arg Tyr Ala Ala Val Leu Leu Arg Pro Phe Leu Thr His Ala Pro Lys
157          465          470          475          480
158 Glu Ile Phe Glu Gln Leu Asn Ile Asn Asn Pro Gln Phe Met Glu Phe
159          485          490          495
160 Ser Ser Leu Glu Gln Tyr Gly Val Leu Thr Glu Ser Ile Met Val Thr
161          500          505          510
162 Gly Gln Pro Lys Pro Ile Phe Pro Arg Leu Asp Ser Glu Ala Glu Ile
163          515          520          525
164 Ala Tyr Ile Lys Glu Ser Met Gln Pro Pro Ala Thr Glu Glu Glu Lys
165          530          535          540
166 Glu Glu Ile Pro Ser Lys Pro Gln Ile Asp Ile Lys Asp Phe Asp Lys
167          545          550          555          560
168 Val Glu Ile Lys Ala Ala Thr Ile Ile Asp Ala Glu His Val Lys Lys
169          565          570          575
170 Ser Asp Lys Leu Leu Lys Ile Gln Val Asp Leu Asp Ser Glu Gln Arg
171          580          585          590
172 Gln Ile Val Ser Gly Ile Ala Lys Phe Tyr Thr Pro Asp Asp Ile Ile
173          595          600          605
174 Gly Lys Lys Val Ala Val Val Thr Asn Leu Lys Pro Ala Lys Leu Met
175          610          615          620
176 Gly Gln Lys Ser Glu Gly Met Ile Leu Ser Ala Glu Lys Asp Gly Val
177          625          630          635          640
178 Leu Thr Leu Val Ser Leu Pro Ser Ala Ile Pro Asn Gly Ala Val Ile
179          645          650          655
180 Lys
181
182
183

```